

# BUFFON'S NEEDLE LANDING NEAR BESICOVITCH IRREGULAR SELF-SIMILAR SETS

MATTHEW BOND AND ALEXANDER VOLBERG

**ABSTRACT.** In this paper we get an upper estimate of the Favard length of an arbitrary neighborhood of an arbitrary self-similar Cantor set. Consider  $L$  closed discs of radius  $1/L$  inside the unit disc. By using linear maps of the disc onto the smaller discs we can generate a self-similar Cantor set  $\mathcal{G}$ . One such process is to let  $\mathcal{G}_n$  be the union of all possible images of the unit disc under  $n$ -fold compositions of the similarity maps. Then  $\mathcal{G} = \bigcap_n \mathcal{G}_n$ . One may then ask the rate at which the Favard length – the average over all directions of the length of the orthogonal projection onto a line in that direction – of these sets  $\mathcal{G}_n$  decays to zero as a function of  $n$ . Previous quantitative results for the Favard length problem were obtained by Peres–Solomyak [19] and Tao [21]; in the latter paper a general way of making a quantitative statement from the Besicovitch theorem is considered. But being rather general, this method does not give a good estimate for self-similar structures such as  $\mathcal{G}_n$ . In the present work we prove the estimate  $\text{Fav}(\mathcal{G}_n) \leq e^{-c\sqrt{\log n}}$ . While this estimate is vastly improved compared to [19] and [21], it is worse than the power estimate  $\text{Fav}(\mathcal{G}_n) \leq \frac{C}{n^p}$  proved for specific sets  $\mathcal{G}_n$  with additional product structures in Nazarov-Peres-Volberg [17] and Laba-Zhai [11]. The power estimate still appears to be related to a certain regularity property of zeros of a corresponding linear combination of exponents (we call this property *analytic tiling*). We consider also the Sierpinski gasket, where this regularity of zeros exists, resulting in an improvement to a power estimate.

## 1. Introduction

Let  $E \subset \mathbb{C}$ , and let  $\text{proj}_\theta$  denote orthogonal projection onto the line having angle  $\theta$  with the real axis. The **average projected length** or **Favard length** of  $E$ ,

---

1991 *Mathematics Subject Classification*. Primary: 28A80. Fractals, Secondary: 28A75, Length, area, volume, other geometric measure theory 60D05, Geometric probability, stochastic geometry, random sets 28A78 Hausdorff and packing measures.

Research of the authors was supported in part by NSF grants DMS-0501067, 0758552 .

$\text{Fav}(E)$ , is given by

$$\text{Fav}(E) = \frac{1}{\pi} \int_0^\pi |\text{proj}_\theta(E)| d\theta.$$

For bounded sets, Favard length is also called **Buffon needle probability**, since up to a normalization constant, it is the likelihood that a long needle dropped with independent, uniformly distributed orientation and distance from the origin will intersect the set somewhere.

Consider  $L$  closed discs of radius  $1/L$  inside the unit disc. By using linear maps of the unit disc onto the smaller discs we can generate a self-similar Cantor set  $\mathcal{G}$ . A partial construction  $\mathcal{G}_n$  of  $\mathcal{G}$  consists of the union of all possible images of the unit disc under  $n$ -fold compositions of these similarity maps. Then  $\mathcal{G} = \bigcap_n \mathcal{G}_n$ .

One may then ask the rate at which the Favard length of these sets  $\mathcal{G}_n$  decays to zero as a function of  $n$ <sup>(1)</sup>. Observe that  $\mathcal{G}_n$  is in some sense comparable to an  $L^{-n}$  neighborhood of  $\mathcal{G}$ <sup>(2)</sup>, so  $\text{Fav}(\mathcal{G}_n)$  is comparable to the likelihood that “Buffon’s needle” will land in a  $L^{-n}$ -neighborhood of  $\mathcal{G}$ .

The first quantitative results for the Favard length problem were obtained in [19],[21]; in the latter paper a general way of making a quantitative statement from the Besicovitch theorem is considered. But being rather general, this method does not give a good estimate for self-similar structures such as  $\mathcal{G}_n$ .

Indeed, vastly improved estimates have been proven in these cases: in [17], it was shown that for  $1/4$  corner Cantor set one has  $p < 1/6$ , such that  $\text{Fav}(\mathcal{K}_n) \leq \frac{c_p}{n^p}$ , and in [5], [6] the same type power estimate was proved for the Sierpinski gasket  $\mathcal{S}_n$  for some other  $p > 0$ . These results cannot possibly be improved to  $p = 1$ :  $\text{Fav}(\mathcal{K}_n) \geq c \frac{\log n}{n}$ . (This is [1]<sup>(3)</sup>, and the argument and result also apply to  $\mathcal{S}_n$ .) Compare this with [19], in which it was shown that certain random sets of which  $\mathcal{K}_n$  is a special case almost surely decay in Favard length like  $\frac{1}{n}$  in the liminf.

Crucial to [1] was a tiling property: namely, under orthogonal projection on the line with slope  $1/2$ , the squares composing  $\mathcal{K}_n$  tile a line segment. Oddly enough, such a property can be used to prove upper bounds as well: under the assumption that some orthogonal projection in some direction contains an interval, Laba and

---

<sup>1</sup>Such decay must occur by the Besicovitch projection theorem and by continuity of measures, since one takes the Lebesgue measure of decreasing sets in the parameter space of  $\{\text{directions}\} \times \{\text{projected x values}\}$ .

<sup>2</sup>an  $L^{-n}$ -neighborhood of  $\mathcal{G}$  is contained in several small translates of  $\mathcal{G}_n$ , while  $\mathcal{G}_n$  is contained in a neighborhood of size  $\approx L^{-n}$  of  $\mathcal{G}$

<sup>3</sup>the method is stable under “bending the needle” slightly - see [7].

Zhai [11] showed that the result of [17] holds for Cantor-like product sets of finite  $H^1$  measure (but with a smaller exponent). Their argument uses tiling results obtained in Kenyon [10] and Lagarias-Wang [12] to fill in a gap where [17] fails to generalize (more on this shortly).

With the exception of [19] and [21], the above papers all extract their results from information about  $L^2$  norms of the projection multiplicity function, which counts how many squares (or discs) project to cover each point. The function  $f_{n,\theta} : \mathbb{R} \rightarrow \mathbb{N}$  is defined by

$$f_{n,\theta} = \sum_{\text{discs } T \text{ of } \mathcal{G}_n} \chi_{\text{proj}_\theta(T)}.$$

Note that  $\text{Fav}(\mathcal{G}_n) = \pi^{-1} \int_0^\pi |\text{supp}(f_{n,\theta})| d\theta$ . In [17] and [5], the  $L^2$  norm of the analog of this function for squares was studied to obtain Buffon needle probability estimates for  $\mathcal{K}_n$  – in [5],  $p = 1, 2$ , were related to  $\chi_{\text{supp}(f_{n,\theta})}$  via the Cauchy inequality, while in [17],  $p = 2$  was studied via Fourier transforms and related to the measure of the level sets of  $f_{n,\theta} = f$ .

Consider some heuristics. Forgetting for the moment about angles, let  $f : [0, 1] \rightarrow \mathbb{N}$  be any sum of measurable characteristic functions such that  $\|f\|_{L^1} = 1$ . If the mass is concentrated on a small set, the  $L^p$  norm should be large for  $p > 1$ . Thus a large  $L^p$  norm should indicate that the support of a function is small, and vice versa. Let  $K > 0$ , let  $A = \text{supp}\{f\}$ , and let  $A_K = \{x : f \geq K\}$ .  $1 = \int f \leq \|f\|_p \|\chi_A\|_q$ , so  $m(A) \geq \|f\|_p^{-q}$ , a decent estimate. The other basic estimate is not so sharp:  $m(A) \leq 1 - (K - 1)m(A_K)$ . However, a combinatorial self-similarity argument of [17] shows that for the Favard length problem, it bootstraps well under further iterations of the similarity maps – this argument is revisited in Section 4. Hence, up to some loss of sharpness, it has been shown that to study Favard length of these self-similar sets, it is sufficient to study the  $L^2$  norms of  $f_{n,\theta}$ .<sup>(4)</sup>

One must average  $|\text{supp } f_{n,\theta}|$  over the parameter  $\theta$  to get Favard length of  $\mathcal{G}_n$ . For  $\mathcal{K}_n$  and  $\mathcal{S}_n$ , there are some directions for which the orthogonal projections do not even decay to length zero with  $n$  (i.e., the  $L^2$  norms of  $f_{n,\theta}$  are bounded for these angles), and this countable dense set of directions is to a large extent classified in [10]. In [17], a method for controlling the measure of a set of angles  $E$  on which the projections fail to decay rapidly was found: one takes the Fourier transform of

---

<sup>4</sup>So far, only  $L^p$  for  $p = 1, 2$ , or  $\infty$  have played any useful role, to our knowledge.

$f_{n,\theta}$  in the length variable, and takes a sample integral of  $|\hat{f}_{n,\theta}(x)|^2$  over a chosen small interval  $I$  where  $\int_{E \times I} |\hat{f}_{n,\theta}(x)|^2 d\theta dx$  is small. One then shows that there is a  $\theta \in E$  such that  $\int_I |\hat{f}_{n,\theta}(x)|^2 dx$  is large relative to  $|E|$ , and so  $|E|$  must be small.

In all cases,  $\hat{f}_{n,\theta}$  is a decay factor times a self-similar product  $\prod_k \varphi_\theta(L^{-k}y)$  of trigonometric polynomials  $\varphi_\theta$ . The danger is that the low-frequency zeroes might kill off the better-behaved high-frequency terms. In [17], the four frequencies of  $\varphi_\theta$  were symmetric around 0, allowing the terms to simplify to two cosines, and trigonometric identities allowed the whole product to be estimated by a single sine term. In [11], an analogous role was played by tilings of the line on the non-Fourier side by  $proj_{\theta_0}(\mathcal{G}_n)$  in the special direction  $\theta_0$ , and the product structure of  $\mathcal{G}_n$  allowed for a change and separation of variables.

Separating variables is more difficult when there is no product structure. The simplest case without the product structure is the Sierpinski gasket  $\mathcal{S}$  considered in Section 2. We give there a sketch of the power estimate (proven in detail in [5]), which is based on the fact that zeroes of  $\varphi(3^k \cdot)$  are separated away from each other for different values of  $k$ . This special structure of zeros (we call it “analytic tiling” after [11]) is not always available for all angles. We have not yet found an adequate substitute for it in the general case, and this is why for the general case we still only have  $Fav(\mathcal{G}_n) \leq e^{-\epsilon_0 \sqrt{\log n}}$ .

Rather strangely, a claim in the spirit of the Carleson Embedding Theorem, in the form of Lemma 22, plays an important part in our reasoning. Because the Fourier transform turns stacks of discs (i.e., sums of overlapping characteristic functions) into clusters of frequencies, this lemma provides important upper bounds when  $\theta$  belongs to  $E$ .

The main result of this article is the following estimate.

**Theorem 1.** *For all  $n \in \mathbb{N}$ ,*

$$Fav(\mathcal{G}_n) \leq C e^{-\epsilon_0 \sqrt{\log n}}.$$

For the Sierpinski gasket, the result is exactly that of [17]:

**Theorem 2.** *For all  $p < 1/12$ , there exists  $C_p > 0$  such that for all  $n \in \mathbb{N}$ ,*

$$Fav(\mathcal{S}_n) \leq C_p n^{-p}.$$

**Acknowledgements.** We are greatly indebted to Fedja Nazarov for many valuable discussions, we also express our deep gratitude to Izabella Laba for many

useful conversations and to John Garnett who introduced the subject to one of the authors.

## 2. DEFINITIONS AND RESULT FOR SIERPINSKI GASKET. SKETCH OF THE PROOF FOR SIERPINSKI GASKET

We give in the first section the sketch of the estimate for a special self-similar set—the Sierpinski gasket. The structure of zeros of a certain trigonometric polynomial related to this set plays the crucial part in this “good” (meaning power) estimate of the type  $\leq n^{-c}$ . We put this sketch here for the reader to be able to compare it with the general case, where we get only the estimate  $\leq e^{-c\sqrt{\log n}}$  due to the lack of understanding of these zeros. It is elaborated in more detail in [5]. It may be instructive to compare the general and the special cases. If the reader is interested only in the general case then he/she can skip the present section and go directly to the next one. Conversely, the reduced detail and difficulty may benefit the first-time reader looking for the general overview of the method.

$B(z_0, r) := \{z \in \mathbb{C} : |z - z_0| < r\}$ . For  $\alpha \in \{-1, 0, 1\}^n$  let

$$z_\alpha := \sum_{k=1}^n \left(\frac{1}{3}\right)^k e^{i\pi[\frac{1}{2} + \frac{2}{3}\alpha_k]}, \quad \mathcal{S}_n := \bigcup_{\alpha \in \{-1, 0, 1\}^n} B(z_\alpha, 3^{-n}).$$

This set is our approximation of a partial Sierpinski gasket; it is strictly larger. We may still speak of the approximating discs as “Sierpinski triangles.”

The result for the Sierpinski gasket is the following:

**Theorem 3.** *For some  $c > 0$ ,  $\text{Fav}(\mathcal{S}_n) \leq \frac{C}{n^c}$ .*

We will simplify the proof by picking specific values for constants; at the end of this paper, a short remark shows how to recover the full range  $c < 1/6$  as in Theorem 2.

The set  $\mathcal{S}_n$  is  $3^{-n}$  approximation to a Besicovitch irregular set (see [8] for definition) called the Sierpinski gasket. Recently one detects a considerable interest in estimating the Favard length of such  $\epsilon$ -neighborhoods of Besicovitch irregular sets, see [19], [21], [17], [11]. In [19] a random model of such Cantor set is considered and estimate  $\asymp \frac{1}{n}$  is proved. But for non-random self-similar sets the estimates of [19] are more in terms of  $\frac{1}{\log \dots \log n}$  (number of logarithms depending on  $n$ ) and

more suitable for general class of “quantitatively Besicovitch irregular sets” treated in [21].

As in the introduction, let

$$f_{n,\theta} := \sum_{\text{Discs } D \text{ of } \mathcal{S}_n} \chi_{\text{proj}_\theta(D)}.$$

Self-similarity allows us to write  $f_{n,\theta}$  in a form well-suited to Fourier analysis:

$f_{n,\theta} = \frac{1}{2}\nu_n * 3^n \chi_{[-3^{-n}, 3^{-n}]}$ , where

$$\nu_n := *_{k=1}^n \tilde{\nu}_k \text{ and } \tilde{\nu}_k := \frac{1}{3} [\delta_{3^{-k} \cos(\pi/2-\theta)} + \delta_{3^{-k} \cos(-\pi/6-\theta)} + \delta_{3^{-k} \cos(7\pi/6-\theta)}].$$

For  $K > 0$ , let  $A_K := A_{K,n,\theta} := \{x : f_{n,\theta} \geq K\}$ . Let  $\mathcal{L}_{\theta,n} := \text{proj}_\theta(\mathcal{G}_n) = A_{1,n,\theta}$ . For our result, some maximal versions of these are needed:

$$f_{N,\theta}^* := \max_{n \leq N} f_{n,\theta}, \quad A_K^* := A_{K,n,\theta}^* := \{x : f_{n,\theta}^* \geq K\}.$$

Also, let  $E := E_N := \{\theta : |A_K^*| \leq K^{-3}\}$  for  $K = N^{\epsilon_0}$ , where  $\epsilon_0 > 0$  is a small enough absolute constant.

Later, we will jump to the Fourier side, where the function

$$\varphi_\theta(x) := \frac{1}{3} [e^{-i \cos(\pi/2-\theta)x} + e^{-i \cos(-\pi/6-\theta)x} + e^{-i \cos(7\pi/6-\theta)x}]$$

plays the central role:  $\widehat{\nu_n}(x) = \prod_{k=1}^n \varphi_\theta(3^{-k}x)$ .

**2.1. General philosophy.** Fix  $\theta$ . If the mass of  $f_{n,\theta}$  is concentrated on a small set, then  $\|f_{n,\theta}\|_p$  should be large for  $p > 1$  - and vice versa.  $1 = \int f \leq \|f_{n,\theta}\|_p \|\chi_{\mathcal{L}_{\theta,n}}\|_q$ , so  $m(\mathcal{L}_{\theta,n}) \geq \|f\|_p^{-q}$ , a decent estimate. The other basic estimate is not so sharp:

$$m(\mathcal{L}_{\theta,N}) \leq 1 - (K-1)m(A_{K,N,\theta}) \quad (2.1)$$

However, a combinatorial self-similarity argument of [17] and revisited in [1] shows that for the Favard length problem, it bootstraps well under further iterations of the similarity maps:

**Theorem 4.** *If  $\theta \notin E_N$ , then  $|\mathcal{L}_{\theta,NK^3}| \leq \frac{C}{K}$ .*

This is proved in full detail as Theorem 17. Note that the maximal version  $f_N^*$  is used here. A stack of  $K$  triangles at stage  $n$  generally accounts for more stacking per step the smaller  $n$  is. Thus the maximal function  $f_N^*$  captures this information by recording a large level set at height  $K$  whenever  $f_n$  attains height  $K$  for a small value of  $n$ . For fixed  $x \in A_{K,N,\theta}^*$ , the above theorem considers the smallest  $n$  such that  $x \in A_{K,n,\theta}$ , and uses self-similarity and the Hardy-Littlewood

theorem to prove its claim by successively refining an estimate in the spirit of (2.1). Of course, now Theorem 3 follows from the following:

**Theorem 5.** *Let  $\epsilon_0 < 1/\log_3(169)$ , sufficiently,  $\epsilon_0 \leq 1/9.262$ . Then for  $N \gg 1$ ,  $|E| < N^{-\epsilon_0}$ .*

It turns out that  $L^2$  theory on the Fourier side is of great use here. The following is later proved as Theorem 26:

**Theorem 6.** *For all  $\theta \in E_N$  and for all  $n \leq N$ ,  $\|f_{n,\theta}\|_{L^2}^2 \leq CK$ .*

One can then take small sample integrals on the Fourier side and look for lower bounds as well. Let  $K = N^{\epsilon_0}$ , and let  $m = 2\epsilon_0 \log_3 N$ . Theorem 6 easily implies the existence of  $\tilde{E} \subset E$  such that  $|\tilde{E}| > |E|/2$  and number  $n$ ,  $N/4 < n < N/2$ , such that for all  $\theta \in \tilde{E}$ ,

$$\int_{3^{n-m}}^{3^n} \prod_{k=0}^n |\varphi_\theta(3^{-k}x)|^2 dx \leq \frac{2CKm}{N} \leq 2\epsilon_0 N^{\epsilon_0-1} \log N.$$

Number  $n$  does not depend on  $\theta$ ;  $n$  can be chosen to satisfy the estimate in the average over  $\theta \in E$ , and then one chooses  $\tilde{E}$ . Let  $I := [3^{n-m}, 3^n]$ .

Now the main result amounts to this (with absolute constant  $\alpha$  large enough):

**Theorem 7.**

$$\exists \theta \in \tilde{E} : \int_I \prod_{k=0}^n |\varphi_\theta(3^{-k}x)|^2 dx \geq c 3^{m-2\alpha m} = c N^{-2\epsilon_0(2\alpha-1)}.$$

The result:  $2\epsilon_0 \log N \geq N^{1-\epsilon_0(4\alpha-1)}$ , i.e.,  $N \leq N^*$ . Now we sketch the proof of Theorem 7. We split up the product into two parts: high and low-frequency:  $P_{1,\theta}(z) = \prod_{k=0}^{n-m-1} \varphi_\theta(3^{-k}z)$ ,  $P_{2,\theta}(z) = \prod_{k=n-m}^n \varphi_\theta(3^{-k}z)$ .

**Theorem 8.** *For all  $\theta \in E$ ,  $\int_I |P_{1,\theta}|^2 dx \geq C 3^m$ .*

Low frequency terms do not have as much regularity, so we must control the damage caused by the **set of small values**,  $SSV(\theta) := \{x \in I : |P_2(x)| \leq 3^{-\ell}\}$ ,  $\ell = \alpha m$ . In the next result we claim the existence of  $\mathcal{E} \subset \tilde{E}$ ,  $|\mathcal{E}| > |\tilde{E}|/2$  with the following property:

**Theorem 9.**

$$\int_{\tilde{E}} \int_{SSV(\theta)} |P_{1,\theta}(x)|^2 dx d\theta \leq 3^{2m-\ell/2} \Rightarrow \forall \theta \in \mathcal{E} \int_{SSV(\theta)} |P_{1,\theta}(x)|^2 dx \leq c K 3^{2m-\ell/2}.$$

Then Theorems 8 and 9 give Theorem 7; since  $\ell = \alpha m$  and  $K^2 = 3^m$ , we see that any  $\alpha > 2$  may be used for this estimate; however, we will need  $\alpha$  to be larger soon.

**2.2. Locating zeros of  $P_2$ .** We can consider  $\Phi(x, y) = 1 + e^{ix} + e^{iy}$ . The key observations (see also the discussion section at the end of the paper) are

$$|\Phi(x, y)|^2 \geq a(|4 \cos^2 x - 1|^2 + |4 \cos^2 y - 1|^2), \quad \frac{\sin 3x}{\sin x} = 4 \cos^2 x - 1. \quad (2.2)$$

Actually, we will set  $\alpha = a^{-1}$  in the end. Changing variable we can replace  $3\varphi_\theta(x)$  by  $\phi_t(x) = \Phi(x, tx)$ .

Consider  $P_{2,t}(x) := \prod_{k=n-m}^n \frac{1}{3} \phi_t(3^{-k}x)$ ,  $P_{1,t}(x) := \prod_{k=0}^{n-m} \frac{1}{3} \phi_t(3^{-k}x)$ .

We need  $SSV(t) := \{x \in I : |P_{2,t}(x)| \leq 3^{-\ell}\}$ . One can easily imagine it if one considers  $\Omega := \{(x, y) \in [0, 2\pi]^2 : |\mathcal{P}(x, y)| := |\prod_{k=0}^m \Phi(3^k x, 3^k y)| \leq 3^{m-\ell}\}$ . Moreover, (using that if  $x \in SSV(t)$  then  $3^{-n}x \geq 3^{-m}$ , and using  $x dx dt = dx dy$ ) we change variable in the next integral:

$$\begin{aligned} \int_{\tilde{E}} \int_{SSV(t)} |P_{1,t}(x)|^2 dx dt &= 3^{-2n+2m} \cdot 3^n \int_{\tilde{E}} \int_{3^{-n}SSV(t)} \left| \prod_{k=m}^n \Phi(3^k x, 3^k tx) \right|^2 dx dt \leq \\ &3^{-n+3m} \int_{\Omega} \left| \prod_{k=m}^n \Phi(3^k x, 3^k y) \right|^2 dx dy. \end{aligned}$$

Now notice that by our key observations

$$\Omega \subset \{(x, y) \in [0, 2\pi]^2 : |\sin 3^{m+1}x|^2 + |\sin 3^{m+1}y|^2 \leq a^{-m} 3^{2m-2\ell} \leq 3^{-\ell}\}. \quad (2.3)$$

The latter set  $\mathcal{Q}$  is the union of  $4 \cdot 3^{2m+2}$  squares  $Q$  of size  $3^{-m-\ell/2} \times 3^{-m-\ell/2}$ . Fix such a  $Q$  and estimate

$$\begin{aligned} \int_Q \left| \prod_{k=m}^n \Phi(3^k x, 3^k y) \right|^2 dx dy &\leq 3^\ell \int_Q \left| \prod_{k=m+\ell/2}^n \Phi(3^k x, 3^k y) \right|^2 dx dy \leq \\ &3^\ell \cdot (3^{-m-\ell/2})^2 \int_{[0, 2\pi]^2} \left| \prod_{k=0}^{n-m-\ell/2} \Phi(3^k x, 3^k y) \right|^2 dx dy \leq 3^\ell \cdot (3^{-m-\ell/2})^2 \cdot 3^{n-m-\ell/2} = 3^{-2m} \cdot 3^{n-m-\ell/2}. \end{aligned}$$

Therefore, taking into account the number of squares  $Q$  in  $\mathcal{Q}$  and the previous estimates we get

$$\int_E \int_{SSV(t)} |P_{1,t}(x)|^2 dx dt \leq 3^{2m-\ell/2}.$$

Theorem 9 is proved.



**Remarks.** The rest of the paper is devoted to general self-similar sets, where we can get only a  $e^{-c\sqrt{\log n}}$  result due to the lack of structure (possible lack of “analytic tiling”) of the zeros of trigonometric polynomials, which are “telescopic products” of one trigonometric polynomial. See the last section of this work for the discussion.

It is true that  $\alpha$  depends on the constant  $a$  in (2.2), since it appears in (2.3). One can use  $a = \frac{1}{18}$ , attained at  $(x, y) = (0, \pi)$ . Then from (2.3), we get  $\alpha = m/\ell \geq \log_3(162) \approx 4.631$  as our last condition on  $\alpha$ . We need this to compute the best exponent  $p$ .

Note that in our argument, we cut a couple corners. To get the best exponent currently available, let  $\gamma > 1$ . Let  $m = \gamma\epsilon_0 \log_3 N$ . Then the argument works as long as  $\epsilon_0 < [2\gamma\alpha + 1 - \gamma]^{-1}$ , i.e.,  $\epsilon_0 < \frac{1}{2\log_3(169)}$ . Using the sharper exponent  $\beta > 2$  in Theorem 17, one can get any  $p = \frac{1}{\epsilon_0^{-1} + \beta} < \frac{1}{[2\log_3(169)]^{-1} + 2}$  in the estimate  $\text{Fav}(\mathcal{S}_n) \leq C_p n^{-p}$ . In particular,  $p < \frac{1}{11.262}$  is small enough.

This can be improved if more care is taken, but not beyond  $p = 1/6$  without a substantial change in the overall approach.

### 3. THE FOURIER-ANALYTIC PART

**3.1. The setup.** The goal of this section is to prove Theorem 11, which shows that for most directions, a considerable amount of stacking occurs when the discs are projected down. Throughout the paper, the constants  $c$  and  $C$  will vary from line to line, but will be absolute constants not depending on anything. The symbols  $c$  and  $C$  will typically denote constants that are sufficiently small or large, respectively. Everywhere we use the definition  $B(z_0, \varepsilon) := \{z \in \mathbb{C} : |z - z_0| < \varepsilon\}$ .

Let

$$\mathcal{G}_1 := \bigcup_{j=1}^L B(r_j e^{i\theta_j}, \frac{1}{L}).$$

Then one constructs  $\mathcal{G}$  and  $\mathcal{G}_n$  using the similarity maps of the unit disc onto the discs forming  $\mathcal{G}_1$ . For convenience, we will now rescale  $\mathcal{G}_n$  by a factor absolutely comparable to 1 and bound the discs of  $\mathcal{G}_n$  by slightly larger discs and study this set instead.

Recall that

$$f_{n,\theta} := \sum_{\text{Discs } D \text{ of } \mathcal{G}_n}^L \chi_{\text{proj}_\theta(D)},$$

Observe that  $f_{n,\theta} = \nu_n * L^n \chi_{[-L^{-n}, L^{-n}]}$ , where  $\nu_n := *_{k=1}^n \tilde{\nu}_k$  and

$$\tilde{\nu}_k = \frac{1}{L} \left[ \sum_{l=1}^L \delta_{L^{-k} r_l \cos(\theta - \theta_l)} \right].$$

We will now slightly modify  $f$  for convenience. Note that

$$\hat{f}_{n,\theta}(x) = L^n \hat{\chi}_{[-L^{-n}, L^{-n}]}(x) \cdot \prod_{k=1}^n \phi_\theta(L^{-k}x),$$

where  $\phi_\theta(x) = \frac{1}{L} [\sum_{l=1}^L e^{-ir_l \cos(\theta_l - \theta)x}]$ . By factoring and changing the variable, we may instead write in place of  $\phi_\theta$  the function

$$\varphi_t(x) = \frac{1}{L} [1 + e^{ix} + e^{itx} + \sum_{l=4}^L e^{a_l x + b_l t x}], \quad t \in [0, 1]. \quad (3.1)$$

We assumed here that  $r_1 = 0$ ,  $r_2 = r_3 = 1$ ,  $\theta_2 = 0$ ,  $\theta_3 = \pi/2$ . We can do this by affine change of variable.

For numbers  $K, N > 0$ , define the following:

$$f_N^*(s) := f_{N,t}^* \sup_{n \leq N} f_{n,t}(s) \quad (3.2)$$

$$A_K^* := A_{K,N,t}^* := \{s : f_N^*(s) \geq K\} \quad (3.3)$$

$$E := \{t : |A_K^*| \leq \frac{1}{K^3}\}. \quad (3.4)$$

$E$  is essentially the set of pathological  $t$  such that  $\|f_{n,t}\|_{L^2(s)}$  is small for all  $n \leq N$ , as in [17]. In fact, we have this result, proved in Section 7:

**Theorem 10.** *Let  $t \in E$ . Then*

$$\max_{0 \leq n \leq N} \|f_{n,t}\|_{L^2(s)}^2 \leq cK.$$

The aim of Section 3 is to prove the following:

**Theorem 11.** *Let  $\epsilon_0$  be a fixed small enough constant. Then for  $N \gg 1$ ,  $|E| < e^{-\epsilon_0(\log N)^{1/2}}$ .*

So let  $K \approx e^{\epsilon_0(\log N)^{1/2}}$ , and suppose  $|E| > \frac{1}{K}$ . We will show that  $N < N^*$ , for some finite constant  $N^* \gg 1$ .

**3.2. Initial reductions.** Because of Theorem 10, we have  $\forall t \in E$ ,

$$K \geq \|f_{N,t}\|_{L^2(s)}^2 \approx \|\widehat{f_{N,t}}\|_{L^2(x)}^2 \geq C \int_1^{L^{N/2}} |\widehat{\nu_N}(x)|^2 dx \quad (3.5)$$

Let  $m \approx (\frac{\epsilon_0}{2} \log N)^{1/2}$ ,  $K \approx \log N$ . Split  $[1, L^{N/2}]$  into  $N/2$  pieces  $[L^k, L^{k+1}]$  and take a sample integral of  $|\widehat{\nu_N}|^2$  on a small block  $I := [L^{n-m}, L^n]$ , with  $n \in [N/4, N/2]$  chosen so that

$$\frac{1}{|E|} \int_E \int_{L^{n-m}}^{L^n} |\widehat{\nu_N}(x)|^2 dx dt \leq CKm/N.$$

This choice is possible by (3.5). Define

$$\tilde{E} := \{t \in E : \int_{L^{n-m}}^{L^n} |\widehat{\nu_N}(x)|^2 dx \leq 2CKm/N\}.$$

It then follows that  $|\tilde{E}| \geq \frac{1}{2K}$ .

Note that  $\widehat{\nu_N}(x) = \prod_{k=1}^N \varphi(L^{-k}x) \approx \prod_{k=1}^n \varphi(L^{-k}x)$  for  $x \in [L^{n-m}, L^n]$ .

So for  $t \in E$ ,

$$\int_{L^{n-m}}^{L^n} \prod_{k=1}^n |\varphi_t(L^{-k}x)|^2 dx \leq \frac{CKm}{N} \leq 2\epsilon_0 N^{\epsilon_0-1} \log N.$$

Recall that  $m \approx (\frac{\epsilon_0}{2} \log N)^{1/2}$ . Later, we will show that  $\exists t \in E$  and absolute constant  $\alpha$  such that

$$\int_{L^{n-m}}^{L^n} \prod_{k=1}^n |\varphi_t(L^{-k}x)|^2 dx \geq cL^{m-2\cdot\alpha m^2} \geq cN^{-\alpha\epsilon_0}. \quad (3.6)$$

The result:  $2\epsilon_0 \log N \geq N^{1-4\alpha\epsilon_0-\epsilon_0}$ , i.e.,  $N \leq N^*$  if  $\epsilon_0$  is small enough. In other words:

**Proposition 12.** *Inequality (3.6) is sufficient to prove Theorem (11). Further, inequality 3.6 can be deduced from Propositions 13 and 14, as will be seen shortly.*

So let us prove inequality (3.6).

First, let us write  $\prod_{k=1}^n \varphi_t(L^{-k}x) = P_t(x) = P_{1,t}(x)P_{2,t}(y)$ , where  $P_2$  is the low frequency part, and  $P_1$  is has medium and high frequencies:

$$P_{1,t}(x) := \prod_{k=1}^{n-m} \varphi_t(L^{-k}x) = \widehat{\nu_{n-m}}(x)$$

$$P_{2,t}(x) = \prod_{k=n-m}^n \varphi_t(L^{-k}x) = \widehat{\nu_m}(L^{m-n}x)$$

We want the following:

**Proposition 13.** *Let  $t \in E$  be fixed. Then  $\int_{L^{n-m}}^{L^n} |P_{1,t}(x)|^2 dx \geq C L^m$ .*

Recall that we defined the set  $\tilde{E}, |\tilde{E}| > |E|/2$ , and we assume that

$$|E| > 1/K. \quad (3.7)$$

Recall that we denoted

$$I = [L^{n-m}, L^n].$$

We also want a proportion of the contribution to the integral separated away from the complex zeroes of  $P_{2,t}$ :

**Proposition 14.** *Let  $SSV(t) := \{x \in I : |P_{2,t}(x)| \leq L^{-\alpha m^2}\}$ . Suppose also that  $E$  is unable to hide, that is (3.7) is valid. Then there exists a subset  $\mathcal{E} \subset \tilde{E}$ ,  $|\mathcal{E}| \geq 1/4K$ , such that for every  $\theta \in \mathcal{E}$  one has*

$$\int_{SSV(t)} |P_{1,t}(x)|^2 dx dt \leq 2c L^m,$$

where  $2c$  is less than the  $C$  from Proposition 13. In particular,

$$\frac{1}{|\tilde{E}|} \int_{\tilde{E}} \int_{SSV(t)} |P_{1,t}(x)|^2 dx dt \leq c L^m,$$

**Remarks.** 1) The set  $SSV(t)$  is so named because it is the **set of small values** of  $P_2$  on  $I$ . Combining this with Proposition 13,

$$\int_{L^{n-m}}^{L^n} |P_{1,t}(x)|^2 |P_{2,t}(x)|^2 dx \geq \int_{I \setminus SSV(t)} |P_{1,t}(x)|^2 \cdot L^{-\alpha m^2} dx \geq c L^{m-2\alpha m^2},$$

which gives (3.6)—exactly what we promised to obtain from Propositions 14, 13.

2) Thus Propositions 13 and 14 suffice to prove Theorem 11, and Proposition 12 has been demonstrated.

3) All this holds if (3.7) holds. But if we have the opposite:

$$|E| \leq 1/K = L^{-\frac{m}{2}} = e^{-C(L)\epsilon_0(\log N)^{1/2}}, \quad (3.8)$$

the main result is automatically proved because we have only a small set of singular directions.

First, let us fix  $t \in E$  and prove Proposition 13.

*Proof.* We are using first Salem's trick on

$$\int_0^{L^n} |P_1(x)|^2 dx :$$

Let  $h(x) := (1 - |x|)\chi_{[-1,1]}(x)$ , and note that  $\hat{h}(\alpha) = C \frac{1 - \cos \alpha}{\alpha^2} > 0$ . Then if we write  $P_1 = L^{m-n-1} \sum_{j=0}^{L^{n-m}} e^{i\alpha_j x}$ , we get

$$\begin{aligned} \int_0^{L^n} |P_1(x)|^2 dx &\geq 2 \int_{-L^n}^{L^n} h(L^{-n}x) |P_1(x)|^2 dx \\ &\geq C(L^{m-n})^2 [L^n \cdot L^{n-m} + \sum_{j \neq k; j,k=1}^{L^{n-m}} L^n \hat{h}(L^n(\alpha_j - \alpha_k))] \geq CL^m. \end{aligned}$$

To show that this is not concentrated on  $[0, L^{n-m}]$ , we will use Theorem 10 and Lemma 22. We get

$$\begin{aligned} \int_0^{L^{n-m}} |P_1(x)|^2 dx &= \int_0^{L^{n-m}} |\widehat{\nu_{n-m}}(x)|^2 dx = L^{2(m-n)} \int_0^{L^{n-m}} \left| \sum_{j=0}^{n-m} e^{i\alpha_j x} \right|^2 dx \\ &\leq CK \leq CL^{\frac{m}{2}}. \end{aligned}$$

□

So now we have Proposition 13. The greater challenge will be Proposition 14.

**3.3. The proof of Proposition 14.** Recall that  $SSV(t) := \{x \in I = [L^{n-m}, L^n] : |P_{2,t}(x)| \leq L^{-\alpha m^2}\}$ .

To get Proposition 14, we will split  $P_{1,t}$  into two parts,  $P_{1,t}^\sharp(x)$  and  $P_{1,t}^b(x)$  corresponding to medium and high frequencies.

A straightforward application of Lemma 22 to high frequency part  $P_{1,t}^\sharp(x)$  will get us part of the way there, see Proposition 16 (for fixed  $t$ , the size of  $SSV(t)$  does not overwhelm the average smallness of  $P_{1,t}^\sharp(x)$ ), and the claim 15 applied to medium frequency term  $P_{1,t}^b(x)$  will further sharpen the final estimate to what we need.

Naturally,  $P_{1,t}^b(x)$  and  $P_{1,t}^\sharp(x)$  are defined as the medium and high frequency parts of  $P_{1,t}(x)$ . Below,  $\ell := \alpha m$ :

$$\begin{aligned} P_{1,t}^b(x) &:= \prod_{k=n-m-\ell}^{n-m-1} \varphi_t(L^{-k}x) = \widehat{\nu_{\ell-1}}(L^{m+\ell-n}x), \\ P_{1,t}^\sharp(x) &:= \prod_{k=1}^{n-m-\ell-1} \varphi_t(L^{-k}x) = \hat{\nu}_{n-m-\ell-1}(x). \end{aligned}$$

Here is the first claim of this subsection

**Proposition 15.** *For all sufficiently small positive numbers  $\tau \leq \tau_0$  and for all sufficiently large  $m$  and  $\ell = \alpha m$  there exists an exceptional set  $H$  of directions  $t$  such that*

$$|H| \leq L^{-\ell/2}, \quad (3.9)$$

$$\forall t \notin H \forall x \in [L^{n-m}, L^n], |P_{1,t}^b(x)| \leq e^{-\tau \ell}. \quad (3.10)$$

*Proof.* Notice that

$$\phi_\theta(r) = \Phi(r \cos \theta, r \sin \theta),$$

where for  $x = (x_1, x_2)$ ,

$$\Phi(x) := \Phi(x_1, x_2) = \frac{1}{L} \sum_{l=1}^L e^{2\pi i \langle a_l, x \rangle}.$$

As some pair of vectors  $a_l - a_1$ ,  $l \in [1, L]$  must span a two-dimensional space, we can assume without the loss of generality (make an affine change of variable) that

$$a_1 = (0, 0), a_2 = (1, 0), a_3 = (0, 1).$$

Then

$$\Phi(x_1, x_2) = \frac{1}{L} (1 + e^{2\pi i x_1} + e^{2\pi i x_2} + \sum_{l=4}^L e^{2\pi i \langle a_l, x \rangle}). \quad (3.11)$$

We make the change of variable  $y = (y_1, y_2) = L^{-(n-m)}x$ . Let  $R_t$  denote the ray  $y_2 = t y_1$ . Then we need to prove that there exists a small set  $H$  of  $t$ 's such that if  $y \in R_t \cap \{y : |y| \in [1, L^m]\}$ ,  $t \notin H$  then

$$|\Phi(y) \cdots \Phi(L^\ell y)| \leq e^{-\tau \ell}. \quad (3.12)$$

We consider only the case  $t \in [0, 1]$ , all our  $y$ 's will be such that  $0 < y_2 \leq y_1$ , and as  $|y| \geq 1$  we have  $y_1 \geq \frac{1}{\sqrt{2}}$ .

It is very difficult if at all possible for function  $\Phi$  to satisfy  $|\Phi(y)| = 1$ . In fact, looking at (3.11) we can see that

$$|\Phi(y)| \leq 1 - b \operatorname{dist}(y, \mathbb{Z}^2) \leq e^{-b \operatorname{dist}(y, \mathbb{Z}^2)}. \quad (3.13)$$

Therefore, we are left to understand that there are few  $t$ 's such that

$$\exists y \in R_t, : y_1 \in [\frac{1}{\sqrt{2}}, L^m] : b \cdot \sum_{k=0}^{\ell} \operatorname{dist}(L^k y, \mathbb{Z}^2) \leq \tau \ell. \quad (3.14)$$

Fix  $y \in R_t$  as above. If (3.14) holds then for 90 per cent of  $k$ 's one has

$$\text{dist}(L^k y, \mathbb{Z}^2) \leq 10\tau \ell. \quad (3.15)$$

Denote  $Z_y := \{k \in [0, \ell] : \text{dist}(L^k y, \mathbb{Z}^2) \leq 10\tau \ell\}$ . We know that

$$|Z_y| \geq 0.9\ell.$$

Let us call **scenario** the collection  $s := \{m_1; k_1, \dots, k_{0.1\ell}\}$ , where  $m_1 = 0, \dots, m; 0 \leq k_1 < \dots < k_{0.1\ell}$ .

Every  $t$  such that there exists  $y$  such that (3.14) holds generates several scenarios according to

$$y_1 \in [L^{m_1-1}, L^{m_1})$$

and according to what is the set  $[0, \ell] \setminus Z_y$ —this is the set  $k_1, \dots, k_{0.1\ell}$  of the scenario.

We will calculate the number of scenarios later. Now let us fix a scenario  $s = \{m_1; k_1, \dots, k_{0.1\ell}\}$ , and let us estimate the measure of the set  $T(s)$ ,  $T(s) := \{t \in (0, 1) : \exists y, y_2 = ty_1, y_1 \in [L^{m_1-1}, L^{m_1}) \text{ such that } [0, \ell] \setminus Z_y = \{k_1, \dots, k_{0.1\ell}\}\}$ . To do that for this fixed scenario we fix a **net**. To explain what is a net we fix

$$a := \left\lceil \frac{\log \frac{100}{\eta}}{\log L} \right\rceil + 1,$$

where  $\eta = C\tau$  and  $C$  is an absolute constant to be chosen soon.

A net is a collection  $N(s) := \{n_1, \dots, n_j\}$ ,  $n_1 < n_2 < \dots$ , where every  $n_i$  is not among  $k_j$  included in the scenario,  $j \geq \frac{3\ell}{4a} + 1$ , and

$$n_{i+1} - n_i \geq 2a.$$

Given a scenario it is always possible to build a net. In fact we just delete from  $[0, \ell]$  the numbers  $k_1, \dots, k_{0.1\ell}$  belonging to the scenario, we are left with at least  $0.9\ell$  numbers. We choose an arithmetic progression with step  $a$  (enumerating them anew first). This arithmetic progression will be long enough, its length  $j \geq \frac{3\ell}{4a}$  because after eliminating  $k_1, \dots, k_{0.1\ell}$  we still have at least  $0.9\ell$  numbers left. We mark the numbers of this progression. Then we put back  $k_1, \dots, k_{0.1\ell}$ . The marked numbers will form our net.

If  $t \in T(s)$  then there exists  $y = (y_1, ty_1)$  as above, in particular,

$$\text{dist}(L^{n_i} y, \mathbb{Z}^2) \leq 10\tau \ell, \quad \forall n_i \in N(s).$$

Let us write that then there exist integers  $p_1 \leq q_1$ :  $|L^{n_1}y_1 - q_1| < 10\tau$ ,  $|L^{n_1}y_2 - p_1| < 10\tau$ , so

$$\left| t - \frac{p_1}{q_1} \right| = \left| \frac{L^{n_1}y_2 - p_1}{L^{n_1}y_1 - q_1} - \frac{p_1}{q_1} \right| = \left| \frac{L^{n_1}y_2 - p_1 + p_1}{L^{n_1}y_1 - q_1 + q_1} - \frac{p_1}{q_1} \right|$$

$$\left| \frac{(L^{n_1}y_2 - p_1 + p_1)q_1 - (L^{n_1}y_1 - q_1 + q_1)p_1}{(L^{n_1}y_1 - q_1 + q_1)q_1} \right| \leq \frac{|L^{n_1}y_2 - p_1||q_1| + |L^{n_1}y_1 - q_1||p_1|}{q_1 - 10\tau} \leq 40\tau \frac{1}{q_1}.$$

As promised we choose  $C$ :  $C = 40$ ,  $\eta := 40\tau$  and we get

$$\exists p_1 \leq q_1 : \left| t - \frac{p_1}{q_1} \right| \leq \eta \frac{1}{q_1}. \quad (3.16)$$

Next we choose integers  $p_2 \leq q_2$ :  $|L^{n_2}y_1 - q_2| < 10\tau$ ,  $|L^{n_2}y_2 - p_2| < 10\tau$  and obtain

$$\exists p_2 \leq q_2 : \left| t - \frac{p_2}{q_2} \right| \leq \eta \frac{1}{q_2}. \quad (3.17)$$

Notice also that because of  $|L^{n_1}y_1 - q_1| < 10\eta$ ,  $|L^{n_2}y_1 - q_2| < 10\eta$ ,  $y_1 \geq 1/\sqrt{2}$ , and smallness of  $\tau$ , and the fact that  $n_2 - n_1 \geq 2a$ , we get

$$\frac{q_2}{q_1} \geq L^a \geq \frac{100}{\eta}. \quad (3.18)$$

We continue in the same vein,  $i = 2, \dots, j-1 \geq \frac{3\ell}{4a}$ :

$$\exists p_i \leq q_i : \left| t - \frac{p_i}{q_i} \right| \leq \eta \frac{1}{q_i}. \quad (3.19)$$

Notice also that because of  $|L^{n_1}y_1 - q_1| < 10\eta$ ,  $|L^{n_2}y_1 - q_2| < 10\eta$ ,  $y_1 \geq 1/\sqrt{2}$ , and smallness of  $\tau$ , and the fact that  $n_2 - n_1 \geq 2a$ , we get

$$\frac{q_{i+1}}{q_i} \geq L^a \geq \frac{100}{\eta}. \quad (3.20)$$

Inequality (3.16) gives that  $|T(s)| \leq \eta$ , inequalities (3.16) and (3.17) in conjunction with (3.18) give  $|T(s)| \leq \left(1 + \frac{1}{100}\right)\eta^2$ , similarly all inequalities (3.19), (3.20) together give

$$|T(s)| \leq (1.01\eta)^{\frac{3\ell}{4a}} \geq e^{0.1\ell} L^{-\frac{3}{4}\ell(1-\epsilon(\eta))}.$$

Here we used of course that  $a := \left\lceil \frac{\log \frac{100}{\eta}}{\log L} \right\rceil + 1$ . Finally, if  $\eta$  is sufficiently small we have

$$|T(s)| \leq L^{-\frac{2}{3}\ell}. \quad (3.21)$$

Let  $\mathcal{S}$  denote the set of all scenarios. Now we want to calculate the number of scenarios. This is easy:

$$\#\mathcal{S} \leq m \cdot \binom{\ell}{0.1\ell} \leq \ell \cdot \left(\frac{10}{9}\right)^{0.9\ell} \cdot 10^{0.1\ell}.$$



We just proved that the measure of the set of all  $t \in (0, 1)$  such that one has  
(3.14)

$$\exists y \in R_t, : y_1 \in [\frac{1}{\sqrt{2}}, L^m] : \sum_{k=0}^{\ell} \text{dist}(L^k y, \mathbb{Z}^2) \leq \tau \ell$$

can be estimated as

$$\leq \ell \cdot \left(\frac{10}{9}\right)^{0.9\ell} \cdot 10^{0.1\ell} \cdot L^{-\frac{2}{3}\ell} \leq L^{-\ell/2}.$$

Proposition 15 is proved. We indeed have very few exceptional directions in the sense that on them  $|P_{1,t}^\flat(x)|$  is not uniformly smaller than  $e^{-\tau\ell}$ .

□

Here is the second claim of the subsection:

**Proposition 16.**

$$t \in E \Rightarrow \int_{SSV(t)} |P_{1,t}^\sharp(x)|^2 dx \leq C'' K L^m.$$

We will see in Section 5 that for each  $t$ ,  $SSV(t)$  is contained in  $C \cdot L^m$  neighborhoods of size  $L^{n-m-\ell}$  around the complex zeroes  $\lambda_j$  of  $P_2$ .

Fix  $t$ . Let

$$I_j = [\lambda_j - L^{n-m-\ell}, \lambda_j + L^{n-m-\ell}], \quad (3.22)$$

$$\text{where } SSV(t) \subseteq \bigcup_j I_j \quad (3.23)$$

Choose  $j$  for which  $\int_{I_j} |P_{1,t}^\sharp(x)|^2 dx$  is maximized. Then

$$\int_{SSV(t)} |P_{1,t}^\sharp(x)|^2 dx \leq C L^m \int_{I_j} |P_{1,t}^\sharp(x)|^2 dx \leq C L^m (L^{\ell+m-n})^2 \int_{I_j} \left| \sum_{k=0}^{n-m-\ell} e^{i\alpha_j x} \right|^2.$$

As  $|I_j| \leq 2 \cdot L^{n-m-\ell}$ , so Lemma 22 and the definition of  $E$  give us Proposition 16.

The estimate for  $t \in \tilde{E} \setminus H$  follows. If  $|E| \geq 1/K$ ,  $K = L^{m/2}$ ,  $|\tilde{E}| > 1/2K$ , and we also just proved that  $|H| \leq L^{-\ell/2}$ ,  $\ell = \alpha m$  with large  $\alpha$ , we have a set  $\mathcal{E} \subset \tilde{E} \setminus H$ ,  $|\mathcal{E}| > 1/4K$ , such that for every  $t \in \mathcal{E}$

$$\int_{SSV(t)} |P_1(r)|^2 dr \leq L^{-\ell} \int_{SSV(t)} |P_{1,t}^\sharp(x)(r)|^2 dr \leq C'' K L^m \cdot L^{-\alpha m}.$$

So we proved

$$\int_{SSV(t)} |P_1(r)|^2 dr \leq c L^m \quad (3.24)$$

with  $c$  as small as we wish. In particular, Proposition 14 is completely proved.

## 4. COMBINATORIAL PART

In this section, we show how Theorem 1 follows from Theorem 11.

First, let us define

$$\mathcal{L}_{\theta,N} := \text{proj}_{\theta} \mathcal{G}_N. \quad (4.1)$$

**Theorem 17.** *Let  $\beta > 2$ . (We used  $\beta = 3$  in the previous section). If  $t \notin E$  (see definition (3.4)), then  $|\mathcal{L}_{\theta,NK\beta}| \leq \frac{C}{K}$ .*

*Proof.* Let us use  $\theta$  instead of  $t$  and  $x$  for the space variable on the non-Fourier side, since we do not use Fourier analysis in this proof. Fix  $\theta$  and let  $F := A_K^* = \{x : f_N^*(x) \geq K\}$ . We denote by  $N_x$  the line orthogonal to direction  $\theta$  and passing through  $x$ . We can call it needle at  $x$ . For every  $x \in F$  there are at least  $K$  discs of size  $L^{-r}$ ,  $r = r(x)$ ,  $r \leq N$ , intersecting  $N_x$ . Mark them. Run over all  $x \in F$ . Consider all marked discs. Consider all  $L^{-N}$ -discs that are sub-discs of marked ones. Call them “green”. Let  $U$  be a family of green discs.

We want to show

$$\text{card } U \geq c \cdot K |F| L^N, \quad (4.2)$$

$$|\text{proj}(\cup_{q \in U} q)| \leq \frac{C}{K} \text{card } U L^{-N}, \quad (4.3)$$

Let  $\phi := \sum_{q \in U} \chi_q$ . Then

$$\int \phi dx = \text{card } U L^{-N}.$$

Let  $M$  denote uncentered maximal function. To prove (4.3) it is enough to show that

$$q \in U \Rightarrow \text{proj } q \subset \{x : M\phi(x) > \frac{K}{C}\},$$

and then to use Hardy–Littlewood maximal theorem. But to prove this claim is easy. In fact, let  $x \in \text{proj } q$ ,  $q \in U$ , then there exists  $Q$ —the maximal (by inclusion) marked disc containing  $q$ . Consider  $I := [x - 10\ell(Q), x + 10\ell(Q)]$ . This segment contains the projections of at least  $K$  disjoint discs  $Q_1 := Q, Q_2, \dots, Q_K, \dots$ , of the same sidelength, which intersect  $N_{x_0}$ , where  $x_0$  is a point because of which  $Q = Q_1$  was marked. (The reader should see that  $x_0$  lies really well inside  $I$ .) So  $I$  contains the projections of at least  $\frac{\ell(Q)}{\ell(q)} \cdot K$  green triangles. Whence,

$$\int_I \phi dx \geq \ell(q) \cdot \frac{\ell(Q)}{\ell(q)} \cdot K \geq \frac{1}{20} |I| K.$$

So

$$M\phi(x) > \frac{1}{20} K.$$

We proved (4.3).

Also we proved that  $F \subset \{x : M\phi(x) \geq \frac{K}{20}\}$ . Therefore, by Hardy–Littlewood maximal theorem

$$|F| \leq |\{x : M\phi(x) \geq \frac{K}{20}\}| \leq \frac{C \int \phi}{K} = C \text{card } U L^{-N} K^{-1}.$$

This is (4.2).

Let us estimate  $|\mathcal{L}_{\theta,N} K^\alpha|$  using (4.2) and (4.3). The first step:

$$\begin{aligned} |\mathcal{L}_{\theta,N}| &\leq |\text{proj}(\cup_{q \in U} q)| + L^{-N}(L^N - \text{card } U) \leq \\ &\frac{C}{K} \text{card } U L^{-N} + (L^N - \text{card } U) L^{-N}. \end{aligned}$$

We do not touch the first term, but we improve the second term by using self-similar structure and going to step  $2N$  (inside triangles which are not green there are “green” discs of size  $L^{-2N}$ ). They are just self-similar copies of the original green discs. Then we have the second step:

$$\begin{aligned} |\mathcal{L}_{\theta,N}| &\leq \frac{C}{K} \text{card } U L^{-N} + \text{the rest} \leq \\ &\frac{C}{K} \text{card } U L^{-N} + (L^N - \text{card } U) \frac{C}{K} \text{card } U L^{-2N} + (L^N - \text{card } U)^2 L^{-2N}. \end{aligned}$$

Now we leave first two terms alone and having  $(L^N - \text{card } U)^2$  triangles of size  $L^{-2N}$  we find again “green” discs inside each of those, now green triangles of size  $L^{-3N}$ . They are just self-similar copies of original green discs.

Then we have the third step:

$$\begin{aligned} |\mathcal{L}_{\theta,3N}| &\leq \frac{C}{K} \text{card } U L^{-N} + (L^N - \text{card } U) \frac{C}{K} \text{card } U L^{-2N} + \text{the rest} \leq \\ &\frac{C}{K} \text{card } U L^{-N} + (L^N - \text{card } U) \frac{C}{K} \text{card } U L^{-2N} + (L^N - \text{card } U)^2 \frac{C}{K} \text{card } U L^{-2N} + \\ &\quad (L^N - \text{card } U)^3 L^{-3N}. \end{aligned}$$

After the  $l$ -th step:

$$\begin{aligned} |\mathcal{L}_{\theta,lN}| &\leq \frac{C}{K} \text{card } U L^{-N} (1 + (L^N - \text{card } U) L^{-N} + \dots \\ &\quad + (L^N - \text{card } U)^{l-1} L^{-(l-1)N}) + (L^N - \text{card } U)^l L^{-lN}. \end{aligned}$$

So

$$|\mathcal{L}_{\theta, lN}| \leq \frac{C}{K} \text{card } U L^{-N} \frac{(1 - (1 - \frac{\text{card } U}{L^N})^l)}{(1 - (1 - \frac{\text{card } U}{L^N}))} + e^{-\frac{\text{card } U}{L^N} l} =: I + II.$$

Notice that by (4.2)  $II \leq e^{-K|F|l} \leq e^{-K}$  if the step  $l$  is chosen to be  $l = 1/|F| \leq K^\beta$ . However, we always have  $I \leq \frac{C}{K}$ . So Theorem 17 is completely proved.  $\square$

From Theorems 11 and 17, it is not hard to get Theorem 1.

## 5. PUTTING $SSV(t)$ INTO A FIXED NUMBER OF INTERVALS OF CORRECT SIZE

Now we have to consider  $P_{2,t}(r) = \phi_t(r)\phi_t(L^{-1}r) \cdots \phi_t(L^{-m}r)$ . We are interested in the set

$$SSV(t) := \{r \in [1, L^m] : |P_{2,t}(r)| \leq L^{-Am^2}\}.$$

We will be using so-called Turan's lemma:

**Lemma 18.** *Let  $f(x) = \sum_{l=1}^L c_l e^{\lambda_l x}$ , let  $E \subset I$ ,  $I$  being any interval. Then*

$$\sup_I |f(x)| \leq e^{\max |\Re \lambda_n| |I|} \left( \frac{A|I|}{|E|} \right)^L \sup_E |f(x)|.$$

Here  $A$  is an absolute constant.

In this form it is proved by F. Nazarov [16].

Now let us consider any square  $Q = [x' - 1, x' + 1] \times [-1, 1]$ . We call  $\frac{1}{2}Q$  the concentric square of half the size.

**Lemma 19.** *With uniform constant  $C$  depending only on  $L$  one has*

$$\sup_Q |\phi_t(z)| \leq C \sup_{\frac{1}{2}Q} |\phi_t(z)|.$$

*Proof.* Let  $z_0 = x_0 + iy_0$  is a point of maximum in the closure of  $Q$ . We first want to compare  $|f(z_0)|$  and  $|f(x_0)|$ . Consider  $f_{x_0}(y) := \phi_t(x_0 + iy)$ . Notice that uniformly in  $Q$  and  $x_0$

$$|f'_{x_0}(y)| \leq C(L).$$

This means that  $|f_{x_0}(y)| \geq \frac{1}{2}|f_{x_0}(0)|$  on an interval of uniform length  $c(L)$ .

Notice also that the exponents  $\lambda_l(t), l = 1, \dots, L$ , encountered in  $\phi_t$  are all uniformly bounded. Then applying Lemma 18 we get

$$|\phi_t(z_0)| = |f_{x_0}(y_0)| \leq C'(L)|f_{x_0}(0)|.$$

Now consider  $F(x) = \phi_t(x)$ . We want to compare  $F(x_0) = f_{x_0}(0) = \phi_t(x_0)$  with  $\max_{[x' - \frac{1}{2}, x' + \frac{1}{2}]} |F(x)|$ . By Lemma 18 we get again

$$|f_{x_0}(0)| = |F(x_0)| \leq \sup_{[x' - 1, x' + 1]} |F(x)| \leq C''(L) \sup_{[x' - 1/2, x' + 1/2]} |F(x)| \leq C''(L) \sup_{\frac{1}{2}Q} |\phi_t(z)|$$

Combining the last two display inequalities we get Lemma 19 completely proved.  $\square$

**Lemma 20.** *With uniform constant  $C$  depending only on  $L$  (and not on  $m$ ) one has*

$$\sup_Q |\phi_t(L^{-k}z)| \leq C \sup_{\frac{1}{2}Q} |\phi_t(L^{-k}z)|, \quad k = 0, \dots, m.$$

The proof is exactly the same. We just use  $L^{-k}\lambda_l(t), l = 1, \dots, L$ , encountered in  $\phi_t(L^{-k}\cdot)$  are all uniformly bounded.

By complex analysis lemmas from Section 6 we know that Lemma 20 implies that every  $\frac{1}{2}Q$  has at most  $M$  (depending only on  $L$ ) zeros of  $\phi_t(z)$ . And if we denote them by  $\mu_1, \dots, \mu_M$  then

$$\{x \in \frac{1}{2}Q \cap \mathbb{R} : |\phi_t(x)| \leq L^{-M\ell}\} \subset \cup_{i=1}^M B(\mu_i, L^{-\ell}). \quad (5.1)$$

Consider  $\mu_1, \dots, \mu_S$  being all zeros of  $P_{2,t}$  in  $[1/2, L^m + 1] \times [1/2, 1/2]$ . By abovementioned lemmas from Section 6 and by Lemma 20 we get that

$$S \leq M(L) L^m.$$

From (5.1) it is immediate that

$$\{x \in [1, L^m] : |P_{2,t}(L^{-(n-m)}x)| = |\phi_t(x) \cdots \phi_t(L^{-m}(x))| \leq L^{-M\ell m}\} \subset \cup_{i=1}^{ML^m} B(\mu_i, L^{-\ell}). \quad (5.2)$$

Changing the variable  $y = L^{n-m}x$  we get the structure of the set of small values used above during the proof of Proposition 16:

$$SSV(t) \subset \cup_{i=1}^{CL^m} I_i, \quad (5.3)$$

where each interval  $I_i$  has the length  $2L^{n-m-\ell}$ .

## 6. SOME IMPORTANT STANDARD LEMMAS. A BIT OF COMPLEX ANALYSIS

There are a few important lemmas which we have appealed to repeatedly. The first claim, Lemma 21, uses the Carleson imbedding theorem. A stronger version, Lemma 22, uses general  $H^2$  theory. Its importance lies in its ability to establish a key relationship between the level sets of  $f_{n,t}$  and the  $L^2$  norm of  $\widehat{f_{n,t}}$ . This is because the Fourier transform changes the centers of intervals into the frequencies of an exponential polynomial.

The second claim we split into Lemmas 24 and 25. Given a bounded holomorphic function on the disc, its supremum, and an interior non-zero value, these lemmas bound the number of zeroes and contain the set of small values within certain neighborhoods of these zeroes.

### 6.1. In the spirit of the Carleson imbedding theorem.

**Lemma 21.** *Let  $j = 1, 2, \dots, k$ ,  $c_j \in \mathbb{C}$ ,  $|c_j| = 1$ , and  $\alpha_j \in \mathbb{R}$ . Let  $A := \{\alpha_j\}_{j=1}^k$ . Then*

$$\int_0^1 \left| \sum_{j=1}^k c_j e^{i\alpha_j y} \right|^2 dy \leq C k \cdot \sup_{I \text{ a unit interval}} \# \{A \cap I\}.$$

*Proof.* Let  $A_1 := \{\mu = \alpha + i : \alpha \in A\}$ . Let  $\nu := \sum_{\mu \in A_1} \delta_\mu$ . This is a measure in  $\mathbb{C}_+$ . Obviously its Carleson constant

$$\|\nu\|_C := \sup_{J \subset \mathbb{R}, J \text{ is an interval}} \frac{\nu(J \times [0, |J|])}{|J|}$$

can be estimated as follows

$$\|\nu\|_C \leq 2 \sup_{I \text{ a unit interval}} \# \{A \cap I\}. \quad (6.1)$$

Recall that

$$\forall f \in H^2(\mathbb{C}_+) \quad \int_{\mathbb{C}_+} |f(z)|^2 d\nu(z) \leq C_0 \|\nu\|_C \|f\|_{H^2}^2, \quad (6.2)$$

where  $C_0$  is an absolute constant. Now we compute

$$\begin{aligned} \int_0^1 \left| \sum_{j=1}^k c_j e^{i\alpha_j y} \right|^2 dy &\leq e^2 \int_0^1 \left| \sum_{j=1}^k c_j e^{i(\alpha_j + i)y} \right|^2 dy \leq \\ e^2 \int_0^\infty \left| \sum_{j=1}^k c_j e^{i(\alpha_j + i)y} \right|^2 dy &= e^2 \int_{\mathbb{R}} \left| \sum_{\mu \in A_1} \frac{c_\mu}{x - \mu} \right|^2, \end{aligned}$$

where  $c_\mu := c_j$  for  $\mu = \alpha_j + i$ . The last equality is by Plancherel's theorem.

We continue

$$\begin{aligned} \int_{\mathbb{R}} \left| \sum_{\mu \in A_1} \frac{c_\mu}{x - \mu} \right|^2 &= \sup_{f \in H^2(C_+), \|f\|_2 \leq 1} \left| \left\langle f, \sum_{\mu \in A_1} \frac{c_\mu}{x - \mu} \right\rangle \right|^2 = \\ 4\pi^2 \sup_{f \in H^2(C_+), \|f\|_2 \leq 1} \left| \sum_{\mu \in A_1} c_\mu f(\mu) \right|^2 &\leq C \# \{A_1\} \sup_{f \in H^2(C_+), \|f\|_2 \leq 1} \sum_{\mu \in A_1} |f(\mu)|^2 \leq \\ C \# \{A\} \sup_{f \in H^2(C_+), \|f\|_2 \leq 1} \int_{C_+} |f(z)|^2 d\nu(z) &\leq 2C_0 C \# \{A\} \sup_{I \text{ a unit interval}} \# \{A \cap I\}. \end{aligned}$$

This is by (6.7) and (6.1). The lemma is proved.  $\square$

Now we are going to prove a stronger assertion by a simpler approach. This stronger assertion is what is used in the main part of the article.

**Lemma 22.** *Let  $j = 1, 2, \dots, k$ ,  $c_j \in \mathbb{C}$ ,  $|c_j| = 1$ , and  $\alpha_j \in \mathbb{R}$ . Let  $A := \{\alpha_j\}_{j=1}^k$ . Then Suppose*

$$\int_{\mathbb{R}} \left( \sum_{\alpha \in A} \chi_{[\alpha-1, \alpha+1]}(x) \right)^2 dx \leq S, \quad (6.3)$$

*Then there exists an absolute constant  $C$*

$$\int_0^1 \left| \sum_{\alpha \in A} c_\alpha e^{i\alpha y} \right|^2 dy \leq C S. \quad (6.4)$$

Of course, one can change variables and get:

**Corollary 23.** *Let  $j = 1, 2, \dots, k$ ,  $c_j \in \mathbb{C}$ ,  $|c_j| = 1$ , and  $\alpha_j \in \mathbb{R}$ . Let  $A := \{\alpha_j\}_{j=1}^k$ , and let  $\delta > 0$ . Suppose*

$$\int_{\mathbb{R}} \left( \sum_{\alpha \in A} \chi_{[\alpha-\delta, \alpha+\delta]}(x) \right)^2 dx \leq S, \quad (6.5)$$

*Then there exists an absolute constant  $C$*

$$\int_a^{a+\delta^{-1}} \left| \sum_{\alpha \in A} c_\alpha e^{i\alpha y} \right|^2 dy \leq C S / \delta^2. \quad (6.6)$$

**Remark.** Lemma 22 is obviously stronger than Lemma 21. In fact, let  $S_0$  be the maximal number of points  $A$  in any unit interval. Then

$$f(x) := \sum_{\alpha \in A} \chi_{[\alpha-1, \alpha+1]}(x) \leq 2S_0.$$

Now  $\int_{\mathbb{R}} f^2(x) dx \leq 4kS_0$ , where  $k$  as above is the cardinality of  $A$ . We can put now  $S := 4kS_0$ , apply Lemma 22 and get the conclusion of Lemma 21. The proof of Lemma 22 does not require the Carleson imbedding theorem. Here it is.

*Proof.* Using Plancherel's theorem we write

$$\begin{aligned} \int_0^1 \left| \sum_{\alpha \in A} c_{\alpha} e^{i\alpha y} dy \right|^2 &\leq e \int_0^1 \left| \sum_{\alpha \in A} c_{\alpha} e^{i(\alpha+i)y} dy \right|^2 \leq e \int_0^{\infty} \left| \sum_{\alpha \in A} c_{\alpha} e^{i(\alpha+i)y} dy \right|^2 = \\ &e \int_{\mathbb{R}} \left| \sum_{\alpha \in A} \frac{c_{\alpha}}{\alpha + i - x} \right|^2 dx. \end{aligned}$$

Recall that

$$H^2(\mathbb{C}_+) \text{ is orthogonal to } \overline{H^2(\mathbb{C}_+)} \quad (6.7)$$

Now we continue

$$\begin{aligned} &\int_{\mathbb{R}} \left| \sum_{\alpha \in A} \frac{c_{\alpha}}{\alpha + i - x} \right|^2 dx \leq \\ &\int_{\mathbb{R}} \left| \sum_{\alpha \in A} \frac{c_{\alpha}}{\alpha + i - x} - \sum_{\alpha \in A} \frac{c_{\alpha}}{\alpha - i - x} \right|^2 dx = \\ &\frac{\pi}{2} \int_{\mathbb{R}} \left| \sum_{\alpha \in A} c_{\alpha} P_1(\alpha - x) \right|^2 dx, \end{aligned}$$

where  $P_1$  is the Poisson kernel in the half-plane  $C_+$  at height  $h = 1$ :

$$P_h(x) := \frac{1}{\pi} \frac{h}{h^2 + x^2}.$$

We continue by noticing that  $P_1 * \chi_{[\lambda-1, \lambda+1]}(x) \geq c P_1(\lambda - x)$  with absolute positive  $c$ . This is an elementary calculation, or, if one wishes, Harnack's inequality. Now we can continue

$$\int_0^1 \left| \sum_{\alpha \in A} c_{\alpha} e^{i\alpha y} dy \right|^2 \leq \frac{\pi e}{2c} \int_{\mathbb{R}} \left| (P_1 * \sum_{\alpha \in A} c_{\alpha} \chi_{[\alpha-1, \alpha+1]})(x) \right|^2 dx.$$

Now we use the fact that  $f \rightarrow P_1 * f$  is a contraction in  $L^2(\mathbb{R})$ . So

$$\int_0^1 \left| \sum_{\alpha \in A} c_{\alpha} e^{i\alpha y} dy \right|^2 \leq \frac{\pi e}{2c} \int_{\mathbb{R}} \left| \sum_{\alpha \in A} c_{\alpha} \chi_{[\alpha-1, \alpha+1]}(x) \right|^2 dx \leq C S.$$

The lemma is proved. □



## 6.2. A Blaschke estimate.

**Lemma 24.** *Let  $D$  be the closed unit disc in  $\mathbb{C}$ . Suppose  $\phi$  is holomorphic in an open neighborhood of  $D$ ,  $|\phi(0)| \geq 1$ , and the zeroes of  $\phi$  in  $\frac{1}{2}D$  are given by  $\lambda_1, \lambda_2, \dots, \lambda_M$ . Let  $C = \|\phi\|_{L^\infty(D)}$ . Then  $M \leq \log_2(C)$ .*

*Proof.* Let

$$B(z) = \prod_{k=1}^M \frac{z - \lambda_k}{1 - \bar{\lambda}_k z}.$$

Then  $|B| \leq 1$  on  $D$ , with  $=$  on the boundary. If we let  $g := \frac{\phi}{B}$ , then  $g$  is holomorphic and nonzero on  $\frac{1}{2}D$ , and  $|g(e^{i\theta})| \leq C \forall \theta \in [0, 2\pi]$ . Thus  $|g(0)| \leq C$  by the maximum modulus principle. So we have

$$C \geq |g(0)| = \frac{|\phi(0)|}{|B(0)|} \geq \prod_{k=1}^M \frac{1}{|\lambda_k|} \geq 2^M.$$

□

**Lemma 25.** *In the same setting as Theorem 24, the following is also true for all  $\delta \in (0, 1/3)$ :  $\{z \in \frac{1}{4}D : |\phi| < \delta\} \subseteq \bigcup_{1 \leq k \leq M} B(\lambda_k, \varepsilon)$ , where*

$$\varepsilon := \frac{9}{16}(3\delta)^{1/M} \leq \frac{9}{16}(3\delta)^{1/\log_2(C)}.$$

*Proof.* Let  $\delta \in (0, 1/3)$ , and let  $z \in \frac{1}{4}D$  such that  $|z - \lambda_k| > \varepsilon \forall k$ . Note that  $g$  is harmonic and nonzero on  $\frac{1}{2}D$  with  $|g(0)| \geq 2^M$ . Thus Harnack's inequality ensures that  $|g| \geq \frac{1}{3}2^M$  on  $\frac{1}{4}D$ , so there

$$|\phi(z)| \geq |g(z)B(z)| \geq \frac{1}{3}2^M \prod_{k=1}^M \left| \frac{z - \lambda_k}{1 - \bar{\lambda}_k z} \right| \geq \left( \frac{16\varepsilon}{9} \right)^M \frac{1}{3} = \delta.$$

We can conclude the proof by the contrapositive. □

## 7. COMBINATORIAL THEOREM

For this section, regard the set  $E$  from Section 3 as parameterized by  $\theta$ , and use the variable  $x$  instead of  $s$  on the non-Fourier side, since we will not work on the Fourier side at all during this section.

**Theorem 26.** *Let  $\theta \in E$ . Then*

$$\max_{n: 0 \leq n \leq N} \|f_{n,\theta}\|_{L^2(\mathbb{R})}^2 \leq C K.$$

To prove this we first need the following claim, which is the main combinatorial assertion of this article. It repeats the one in [17] but we give a slightly different proof.

We fix a direction  $\theta$ , we think that the line  $\ell_\theta$  on which we project is  $\mathbb{R}$ . If  $x \in \mathbb{R}$  then by  $N_x$  we denote the line orthogonal to  $\mathbb{R}$  and passing through point  $x$ , we call  $N_x$  a needle. By  $F_L$  we denote  $\{x \in \mathbb{R} : f_N^*(x) := \max_{0 \leq n \leq N} f_{n,\theta}(x) > L\}$  (also known as  $A_L^*$ ).

**Theorem 27.** *There exists an absolute constant  $C$  such that for any large  $K$  and  $M$*

$$|F_{4KM}| \leq C K |F_K| \cdot |F_M|. \quad (7.1)$$

*Proof.* This will be a proof by greedy algorithm. First choose  $y \in F_{4K}$  and consider needle  $N_y$  and discs of certain size  $L^{-j_y}, j_y \leq N$  intersecting  $N_y$ . Consider any family of this sort having more than  $4K$  elements. Fix such a family. We will “fathorize” it, i.e. we consider the father of each element in the family. Two things may happen: 1) there are more than  $4K$  distinct fathers; 2) number of fathers is at most  $4K$ . In the latter case the number of fathers is at least  $2K$ . In fact, we slash the number of elements by fathorizing, but not more than by factor of  $1/2$ . If the first case happens fathorize again, do this till we get to the second case.

After doing this procedure with all  $x \in F_{4K}$  and all families of cardinality bigger than  $4K$  of equal size discs intersecting needle  $N_x$  we come to some awfully complicated set of discs. But we will consider now maximal-by-inclusion discs of this family, the family of these maximal discs is called  $\mathcal{F}_0$ .

Choose disc  $Q_{00} \in \mathcal{F}_0$  such that its sidelength  $\ell(Q_{00})$  is maximal possible in  $\mathcal{F}_0$ . It is very important to notice that  $\mathcal{F}_0$  contains at least  $2K - 1$  discs of the same size as  $Q_{00}$  pierced by a needle  $N_{y_0}$ . This is because of maximality of the lengthsize, the stack pierced by  $N_{y_0}$  could not be eaten up even partially by bigger in size discs from some other stack. So let us call by  $Q_{01}, \dots, Q_{02K-1}, \dots, Q_{0S}, S \geq 2K - 1$ . They are of the same size as  $Q_{00}$  and all intersect a certain needle  $N_{y_0}$ .

Denote

$$I_0 = \text{proj } Q_{00}.$$

Consider all  $q \in \mathcal{F}_0$  such that

$$\text{proj } q \cap 20 I_0 \neq \emptyset.$$

Call them  $\mathcal{F}(Q_{00})$ . Of course  $\ell(q) \leq \ell(Q_{00})$ . For every such  $q$  consider a Cantor square  $Q$ ,  $q \subset Q$ , such that  $\ell(Q) = \ell(Q_{00})$ . Such  $Q$ 's form family  $\tilde{\mathcal{F}}(Q_{00})$ .

**Lemma 28.** *For every  $y \in \mathbb{R}$  the needle  $N_y$  intersects at most  $4K$  discs of the family  $\tilde{\mathcal{F}}(Q_{00})$ .*

*Proof.* Suppose contrary. Then  $N_y$  intersects more than  $4K$  of discs from  $\tilde{\mathcal{F}}(Q_{00})$ . So  $y \in F_{4K}$ , and our pierced family is one of those which we considered at the beginning. It can be fathORIZED. Then the square of size  $\geq 2\ell(Q_{00})$  will be present in  $\mathcal{F}_0$ . Contradiction with maximality of length.  $\square$

**Lemma 29.**  $\text{card } \tilde{\mathcal{F}}(Q_{00}) \leq 88K$ .

*Proof.*

$$\begin{aligned} \text{card } \tilde{\mathcal{F}}(Q_{00}) \cdot \ell(Q_{00}) &= \sum_{Q \in \tilde{\mathcal{F}}(Q_{00})} \ell(Q) \leq \\ &\int_{22I_0} \text{card } \{Q \in \tilde{\mathcal{F}}(Q_{00}) : Q \cap N_y \neq \emptyset\} dy \leq \\ &4K \cdot 22\ell(Q_{00}). \end{aligned}$$

This is by Lemma 28.  $\square$

**Lemma 30.** *There exists an interval  $J_0 \subset I_{y_0}$  such that  $|J_0| \geq c \cdot |I_0|$  with a certain absolute positive  $c$ . And  $J_0 \subset F_K$ .*

*Proof.* We already noticed that  $Q_{00}, Q_{01}, \dots, Q_{02K-1}$  intersect needle  $N_{y_0}$ . Then at least half of them have their center of symmetry to the right of  $N_{y_0}$ , or at least half of them have their center of symmetry to the left of  $N_{y_0}$ . Assume that the first case occurs. Then the segment  $[y_0, c \cdot \ell(Q_{00})]$  obviously is contained in  $F_K$ .  $\square$

**Lemma 31.**  $|F_{4KM} \cap 20I_0| \leq CK \ell(Q_{00}) |F_M| = CK |I_0| |F_M|$ .

*Proof.* Of course  $F_{4KM} \subset F_{4K}$ . For  $y \in F_{4KM} \cap 20I_0$  the whole family of small discs whose quantity is  $> 4KM$  intersecting  $N_y$  will be inside one of those  $Q \in \tilde{\mathcal{F}}(Q_{00})$ , whose number is at most  $88K$  by Lemma 29. Let us enumerate  $Q^1, \dots, Q^s$ ,  $s \leq 88K$  elements of  $\tilde{\mathcal{F}}(Q_{00})$ . So there exists  $i = 1, \dots, s$  such that

$$y \in \text{dilated copy of } F_M \text{ in } \text{proj } Q^i.$$

Hence

$$F_{4KM} \cap 20I_0 \subset \cup_{i=1}^{88K} \text{dilated copy of } F_M \text{ in } \text{proj } Q^i.$$

So

$$|F_{4KM} \cap 20I_0| \leq \sum_{i=1}^{88K} \ell(Q^i) |F_M| \leq 88K \ell(Q_{00}) |F_M|.$$

□

**Lemma 32.**  $|F_{4KM} \cap 20I_0| \leq 88c^{-1}K|F_m| \cdot |J_0|.$

Now we want to repeat all steps for  $F_{4K}^0 := F_{4K} \setminus 20I_0$ . So we fathorize discs pierced by needles  $N_x$ ,  $x \in F_{4K}^0$ . As before we get families  $\mathcal{F}_1$ , maximal sidelength triangle  $Q_{11}$ , families  $\mathcal{F}(Q_{11})$ ,  $\tilde{\mathcal{F}}(Q_{11})$ . Notice that  $\mathcal{F}_1 < \mathcal{F}_0$  in the sense that for every  $q \in \mathcal{F}_1$  there exists  $q \in \mathcal{F}_0$  such that  $q$  is contained in  $Q$ . It is also clear that

$$\ell(Q_{11}) \leq \ell(Q_{00}).$$

Obviously  $Q_{00}, Q_{01}, \dots$  are not in  $\mathcal{F}_1$ , their projections even do not intersect  $\mathbb{R} \setminus 20I_0$ .

There are at least  $2K - 1$  brothers of  $Q_{11}$ :  $Q_{12}, \dots, Q_{12K-1}, \dots$  in  $\mathcal{F}_1$  such that they are of the same size  $\ell(Q_{11})$  and they (and  $Q_{11}$ ) intersect the same needle  $N_{y_1}$ ,  $y_1 \in \mathbb{R} \setminus 20I_0$ . This is again the maximality of the sidelength among  $\mathcal{F}_1$  discs. Let  $I_1 := \text{proj } Q_{11}$ . Notice that

$$I_1 \cap I_0 = \emptyset.$$

In fact,  $y_1 \in I_1, y_1 \notin 20I_0$ ,  $Q_{11}$  size is much smaller than  $20|I_0|$ . We consider all  $q \in \mathcal{F}_1$  such that

$$\text{proj } q \cap (20I_1 \setminus 20I_0) \neq \emptyset.$$

Call this family  $\mathcal{F}(Q_{11})$ . For every  $q \in \mathcal{F}(Q_{11})$  consider Cantor disc  $Q$  containing  $q$  and of the size  $\ell_1 = \ell(Q_{11})$ . Maximal-by-inclusion among such  $Q$ 's form  $\tilde{\mathcal{F}}(Q_{11})$ .

**Lemma 33.** *For any  $y \in \mathbb{R} \setminus 20I_0$ ,  $N_y$  intersects at most  $4K$  discs of  $\tilde{\mathcal{F}}(Q_{11})$ .*

*Proof.* Suppose contrary. Then there exists  $y'_1 \in F_{4K} \cap (\mathbb{R} \setminus 20I_0)$ , and a subfamily of  $\tilde{\mathcal{F}}(Q_{11})$  of cardinality bigger than  $4K$  intersects  $N_{y'_1}$ . It can be fathorized. Then discs of size  $\geq 2\ell(Q_{11})$  would belong to  $\mathcal{F}_1$ . This contradicts the maximality of  $\ell(Q_{11})$ .

□

**Lemma 34.** *For any  $z \in \mathbb{R}$ ,  $N_z$  intersects at most  $8K$  discs of  $\tilde{\mathcal{F}}(Q_{11})$ .*

*Proof.* Suppose contrary. Then there exists  $z \in F_{4K}$ , and a subfamily of  $\tilde{\mathcal{F}}(Q_{11})$  of cardinality bigger than  $4K$  intersects  $N_z$ . Now there is an end-point of  $20I_1 \setminus 20I_0$  (call it  $a$ ), which is closest to  $z$ . Let it be on the right of  $z$ . Then another end-point is also on the right but farther away. As every triangle from the family has a)  $z$  in its projection, and b) a certain point to the right of  $a$  in its projection (their projections intersect  $20I_1 \setminus 20I_0$ —by definition), then all of them have  $a$  in its projection. Let us be lavish and say that 50 percent of them have  $a$  in their projection (the fact is that it is not lavishness, it is necessity: next step will be to consider in the future  $20I_2 \setminus (20I_0 \cup 20I_1)$ , and their can be 2 closest points to  $z$ : one on the left, say,  $b$ , and one on the right, say,  $a$ , and we can guarantee that 50 percent of our discs have either  $b$  or  $a$  in their projections simultaneously). We use the previous Lemma 33, and get that this 5) percent is  $\leq 4K$ . So we are done.  $\square$

**Lemma 35.**  $\text{card } \tilde{\mathcal{F}}(Q_{11}) \leq 172K$ .

*Proof.*

$$\begin{aligned} \text{card } \tilde{\mathcal{F}}(Q_{11}) \cdot \ell(Q_{11}) &= \sum_{Q \in \tilde{\mathcal{F}}(Q_{11})} \ell(Q) \leq \\ &\int_{22I_1} \text{card } \{Q \in \tilde{\mathcal{F}}(Q_{11}) : Q \cap N_y \neq \emptyset\} dy \leq \\ &8K \cdot 22\ell(Q_{11}). \end{aligned}$$

This is by Lemma 28.  $\square$

**Lemma 36.** *There exists an interval  $J_1 \subset I_1$ ,  $|J_1| \leq c \cdot |I_1|$ , such that  $J_1 \subset F_K$ .*

*Proof.* The same proof as for Lemma 30.  $\square$

**Lemma 37.**  $|F_{4KM}^0 \cap 20I_1| \leq C K \ell(Q_{11}) \leq C|, K |I_1|$ .

*Proof.* The same proof as for Lemma 31.  $\square$

Combining Lemmas 36, 37 we get

**Lemma 38.**  $|F_{4KM}^0 \cap 20I_1| \leq C c^{-1} K |J_1|$ .

We continue by introducing

$$F_{4KM}^1 = F_{4KM} \setminus (20I_0 \cup 20I_1).$$

We repeat the whole procedure. There will be  $I_2, J_2 \subset I_2 \cap F_K, |J_2| \geq c \cdot |I_2|$ :

$$I_2 \cap (I_1 \cup I_0) = \emptyset,$$

$$|F_{4KM} \cap 20I_2| \leq Cc^{-1}K|J_2||F_M|,$$

et cetera.

Finally,

$$|F_{4KM}| \leq |F_{4KM} \cap 20I_0| + |(F_{4KM} \setminus 20I_0) \cap 20I_1| + \dots + |(F_{4KM} \setminus 20I_0 \cup 20I_1 \cup \dots \cup 20I_{j-1}) \cap 20I_j| + \dots \leq$$

$$C'K|F_M| \sum_{j=0}^{\infty} |J_j| \leq C'K|F_M||F_K|.$$

We are done with Theorem 27.  $\square$

Now we can prove Theorem 26.

*Proof.* Let  $E_j := \{x : f_{n,\theta}(x) > (4K)^{j+1}\}$ ,  $j = 0, 1, \dots$ . We know by Theorem 27 that

$$|E_j| \leq (CK)^j |E_0|^{j+1}.$$

Hence,

$$\begin{aligned} \int f_{n,\theta}(x)^2 dx &\leq 4K \int f_{n,\theta}(x) dx + \sum_{j=0}^{\infty} \int_{E_j \setminus E_{j+1}} f_{n,\theta}(x)^2 dx \leq \\ &4CK + (4K)^{j+2} (CK)^j |E_0|^{j+1}. \end{aligned}$$

If  $|\{x : f_N^*(x) > K\}| \leq 1/K^{2+\tau}$  then for all  $n \leq N$  we can immediately read the previous inequality as

$$\int f_{n,\theta}(x)^2 dx \leq C(\tau)K.$$

$\square$

## 8. DISCUSSION

**8.1. Difficulties for more general self-similar sets.** The reason we were able to prove the stronger estimate for the Sierpinski gasket is exactly given by (2.2). It is a quantified version of the fact that the three-term sum  $\varphi(z) = 1 + e^{iz} + e^{itz}$  is zero if and only if the summands are  $e^{2j\pi i}$ ,  $j = 0, 1, 2$ , and that for such  $z$ ,  $\varphi(3^k z) = 3$  for all integers  $k \geq 1$ . An alternate argument using this fact in this form is employed in [6]. Both versions of this fact we call by the general term “analytic tiling”.

But there cannot be such a thing in the general case. Suppose we had 5 self-similarities, and that for some direction  $\theta$ , we had  $\phi_\theta(x_0) = 1 + (-i) + i + e^{2\pi i/3} + e^{4\pi i/3} = 0$ . Then clearly, taking fifth powers of the summands results in another zero with exactly the same summands, in complete and utter contrast to the three-point case. Similar examples using partitions into relatively prime roots of unity exist for numbers other than 5.

Though perhaps there is some hope that for arbitrary sets, some other form of analytic tiling occurs for typical directions in the arbitrary case (with small measure of exceptional directions). Ergodic theory may be of central importance. For example, if one considers  $\mathcal{K}_n$  as in [17], one gets  $\varphi_\theta(z) = 1 + e^{i\pi z} + e^{i\lambda z} + e^{i(\lambda+\pi)z}$ , which has the zero  $z = 1$ . Then  $\varphi(4^k) = 2(1 + \cos(4^k\lambda))$  for  $k > 0$ .  $\lambda$  depends continuously on  $\theta$ , and for fixed  $\lambda$  such an ergodic sampling results in a sequence  $a_k := \varphi(4^k)$ , and either:

- 1:  $a_k$  is eventually periodic and non-zero,
- 2:  $a_k$  takes values other than 4 only finitely often,
- or 3 (the case for almost every  $\lambda$ ):  $4^k\lambda \bmod 2\pi$  evenly samples  $[0, 2\pi]$  over the long term, with long-term average  $\frac{1}{N} \sum_{k=1}^N a_k \rightarrow 2$  as  $N \rightarrow \infty$ .

## REFERENCES

1. M. Bateman, A. Volberg, *An estimate from below for the Buffon needle probability of the four-corner Cantor set*, arXiv:math. 0807.2953v1, 2008, pp. 1-11.
2. M. Bateman, N. Katz, *Keakeya sets in Cantor directions*, arXiv:math. 0609187v1, 2006, pp. 1-10.
3. M. Bateman, *Keakeya sets and the directional maximal operators in the plane*, arXiv:math.CA 0703559v1, 2007, pp. 1-20.
4. A. S. Besicovitch, *Tangential properties of sets and arcs of infinite linear measure*, *Bull. Amer. Math. Soc.* **66** (1960), 353-359.
5. M. BOND, A. VOLBERG: *The power law for Buffon's needle landing near the Sierpinski gasket*, arXiv: 0911.0233v2, 2009, pp. 1-35.
6. M. BOND, A. VOLBERG: *Buffon needle lands in  $\epsilon$ -neighborhood of a 1-Dimensional Sierpinski Gasket with probability at most  $|\log \epsilon|^{-c}$* . *Comptes Rendus Mathematique*, Volume 348, Issues 11-12, June 2010, 653-656.
7. M. BOND, A. VOLBERG: *Estimates from below of the Buffon noodle probability for under-cooked noodles*, arXiv:math/0811.1302v1, 2008, pp. 1-10.
8. K. J. Falconer, *The geometry of fractal sets*. Cambridge Tracts in Mathematics, 85. C.U.P., Cambridge-New York, (1986).

9. U. Keich, On  $L^p$  bounds for Kakeya maximal functions and the Minkowski dimension in  $\mathbb{R}^2$ , *Bull. London. Math. Soc.* **31** (1999), pp. 213–221.
10. R. Kenyon, *Projecting the one-dimensional Sierpinski gasket*, *Israel J. Math.* **97** (1997), 221–238.
11. I. Laba, K. Zhai, *Favard length of product Cantor sets*, arXiv:0902.0964v1, Feb. 5 2009.
12. J. C. Lagarias and Y. Wang, *Tiling the line with translates of one tile*, *Invent. Math.* **124** (1996), 341–365.
13. P. Mattila, *Orthogonal projections, Riesz capacities and Minkowski content*, *Indiana Univ. Math. J.* **39** (1990), 185–198.
14. P. Mattila, *Hausdorff dimension, projections, and the Fourier transform*, *Publ. Mat.*, **48** (2004), pp. 3–48.
15. P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces*, Cambridge University Press, 1995.
16. F. Nazarov, *Local estimates of exponential polynomials and their applications to inequalities of uncertainty principle type*, *St Petersburg Math. J.*, v. 5 (1994), No. 4, pp. 3–66.
17. F. Nazarov, Y. Peres, A. Volberg *The power law for the Buffon needle probability of the four-corner Cantor set*, arXiv:0801.2942, 2008, pp. 1–15.
18. Y. Peres, K. Simon and B. Solomyak, *Self-similar sets of zero Hausdorff measure and positive packing measure*, *Israel J. Math.* **117** (2000), 353–379.
19. Y. Peres and B. Solomyak, *How likely is Buffon’s needle to fall near a planar Cantor set?* *Pacific J. Math.* **204**, 2 (2002), 473–496.
20. I. J. Schoenberg, *On the Besicovitch–Perron solution of the Kakeya problem*, *Studies in mathematical analysis and related topics*,
21. T. Tao, *A quantitative version of the Besicovitch projection theorem via multiscale analysis*, pp. 1–28, arXiv:0706.2446v1 [math.CA] 18 Jun 2007.

MATTHEW BOND, DEPT. OF MATH., MICHIGAN STATE UNIVERSITY. [bondmatt@msu.edu](mailto:bondmatt@msu.edu)

ALEXANDER VOLBERG, DEPT. OF MATH., MICHIGAN STATE UNIV. AND THE UNIVERSITY OF EDINBURGH. [volberg@math.msu.edu](mailto:volberg@math.msu.edu)